(12)

LEVEL II

TECHNICAL REPORT 81-2

AN ASSESSMENT OF MARKOV RENEWAL MODELS IN FORECASTING INTERNATIONAL AFFAIRS

DECISION SCIENCE CONSORTIUM, INC.

James O. Chinnis, Jr.
Anthony N.S. Freeling
David A. Seaver

February 1981

AD A097295

DTIC ELECTE APR 0 3 1981 S D F

DTIC FILE COPY

ARPA

Command & Control
Decision & Forecasting
Systems Program

DEFENSE SCIENCES OFFICE • CYBERNETICS TECHNOLOGY DIVISION
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

81 4 2 008

TECHNICAL REPORT 81-2

AN ASSESSMENT OF MARKOV RENEWAL MODELS IN FORECASTING INTERNATIONAL AFFAIRS

by

James O. Chinnis, Jr., Anthony N.S. Freeling, and David A. Seaver

Sponsored by

Defense Advanced Research Projects Agency
Under Contract MDA903-80-C-0194
DARPA Order No. 3831

Under Subcontract from

Decisions and Designs, Inc.
McLean, VA.

February 1981

THE VIEWS AND CONCLUSIONS CONTAINED IN THIS DOCUMENT ARE THOSE OF THE AUTHOR AND SHOULD NOT BE INTERPRETED AS NECESSARILY REPRESENTING THE OFFICIAL POLICIES, EITHER EXPRESSED OR IMPLIED, OF THE DEFENSE ADVANCED RESEARCH PROJECTS AGENCY OR THE UNITED STATES GOVERNMENT.

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

DECISION SCIENCE CONSORTIUM, INC.
7700 Leesburg Pike, Suite 421
Falls Church, Virginia 22043
(703) 790-0510

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

14

1. REPORT NUMBER
TR-81-2

2. GOVT ACCESSION NO.
AD-A097 295

3. RECIPIENT'S CATALOG NUMBER
9

4. TITLE (and Subtitle)
6
An Assessment of Markov Renewal Models in Forecasting International Affairs

5. TYPE OF REPORT & PERIOD COVERED
Technical rept.

6. PERFORMING ORG. REPORT NUMBER
Technical Report 81-2

7. AUTHOR(s)
10
James O. Chinnis, Jr.
Anthony N. S. Freeling
David A. Seaver

8. CONTRACT OR GRANT NUMBER(s)
15
MDA 903-80-C-0194
79-312-0729 (Sub)

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Decision Science Consortium, Inc.
7700 Leesburg Pike, Suite 421
Falls Church, Virginia 22043

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Darpa Order [redacted]-3831

11. CONTROLLING OFFICE NAME AND ADDRESS
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209
11

12. REPORT DATE
Feb[redacted]1981
12

13. NUMBER OF PAGES
24
33

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for publication release: distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Forecasting
Markov renewal models
Bayesian inference

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
See reverse side.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified 390 664

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

A major difficulty for designers of systems to forecast international affairs has been to allow the use of both observable data and subjective estimates. Under DARPA sponsorship, a novel approach has been developed based upon a Bayesian stochastic model. This approach has been developed and demonstrated in a preliminary fashion by other DARPA contractors.

The present report provides a brief appraisal of the approach, offers preliminary suggestions for modifications, and suggests candidate areas of application. The treatment of time and nonstationarity is explored in some detail, and modifications are suggested which could lessen the need to reassess model parameters each time a change occurs in an underlying political process.

SUMMARY

A major difficulty for designers of systems to forecast international affairs has been to allow the use of both observable data and subjective estimates. Under DARPA sponsorship, a novel approach has been developed based upon a Bayesian stochastic model. This approach has been developed and demonstrated in a preliminary fashion by other DARPA contractors.

The present report provides a brief appraisal of the approach, offers preliminary suggestions for modifications, and suggests candidate areas of application. The treatment of time and nonstationarity is explored in some detail, and modifications are suggested which could lessen the need to reassess model parameters each time a change occurs in an underlying political process.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF EXHIBITS

1.0 INTRODUCTION

Among the more interesting and novel forecasting approaches presently being developed by researchers is a Bayesian method for forecasting international affairs based upon the Markov renewal process. At this time, the work in this area has been developed by Duncan and Job under sponsorship of the Cybernetics Technology Division of the Defense Advanced Research Projects Agency (DARPA)(Duncan and Job, 1977; Duncan, 1977; Job, 1977). The present paper is intended as both an evaluative review of the methodology as currently implemented, and as an evaluation of the methodology in a more general sense. The purpose of this evaluation is to advance further this DARPA research area.

2.0 AN EVALUATION OF CURRENT IMPLEMENTATIONS

2.1 Definition of the Markov Renewal Model

Before embarking on a review, a brief outline of the mathematical basis of the model is provided here for readers not familiar with the theory involved.

In general, a Markov renewal process consists of the following elements:

- A set of states, { 1, 2, ... };
- A dimension of time, t;
- A transition matrix, $\underline{\underline{J}}$;
- A matrix of waiting time distributions, $\underline{\underline{F}}$; and
- An initial probability distribution over the states at time 0, $\underline{p}(0)$.

A realization of this process is described by $\{X(t)\}$, where $X(t)$ is the state of the process at time t.

In the application of the Duncan and Job model the state space is finite and small--five being a typical number of states. An element of $\underline{\underline{J}}$, J_{ij}, is taken to be the probability that if a realization is in state i at time t (i.e. $X(t)=i$) then the next transition between states will be to state j. This probability is assumed to be independent of t. Although the general Markov renewal model allows $J_{ii} > 0$, in current applications the concept of a transition into the same state has been considered meaningless, so $J_{ii} = 0$.

An element of $\underline{\underline{F}}$, F_{ij}, is the distribution of the waiting time for the move from state i to state j, conditional upon that being the transition that actually takes place. There are no practical restrictions on these distributions in theory, but in applications of the Duncan and Job models it has been assumed that they are log-normal in form.

In use, the states of the Duncan and Job model are taken to be measures of the stability of the situation, with 1 being the most stable and, typically, 5 being war. The matrices $\underline{\underline{J}}$ and $\underline{\underline{F}}$ are assessed from an analyst, as is the analyst's perception of the current situation, in terms of his view of which state the country is in at the present. With these parameters known, the Markov renewal process is fully specified, and can be used for forecasting. In particular we may calculate $P(X(t)=i)$ for all times t and states i. We may also discover the probabilities that we will have left state i by time t, or have been in state j by time t. Indeed, all statements concerning states and time can, in principle, have probabilities assigned to them. However, in practice, forecasts are taken as $P(X(t)=i)$ for t = 1 day, 2 days, ... , 30 days.

In practical applications of the model, this forecast is redone every day. To accomplish this, the analyst must reassess the current state each day. Duncan and Job have also developed a very elegant Bayesian updating procedure which uses the observed data about changes of state (as assessed by the analyst day-by-day) to update the $\underline{\underline{F}}$ and $\underline{\underline{J}}$ matrices. In this way, observed and subjective data are used together to make the forecasts.

2.2 Performance of Current Implementations

2.2.1 Framework for evaluation. The model being evaluated is a logically consistent one and should be useful in situations in which the following conditions hold:

- the process being forecast behaves approximately as a Markov renewal process; and
- input parameters for the model can be assessed effectively.

Any definitive evaluation of the model would necessarily depend upon extensive empirical testing. In light of the developmental nature of the model at present, such empirical data are not available and the model must be evaluated primarily on rational grounds such as the above.

2.2.2 Appropriateness of the current Markov renewal model. The characterization of international affairs is exceedingly complex. The present model, while requiring extensive prior judgments of model parameters, attempts to simplify the characterization problem by defining a small number of states upon which all other model parameters depend. This limitation of the model's input and output domains to a few states has several implications, as follows:

- Complex observable processes which underlie an analyst's daily assessment of the current state cannot directly influence model behavior; only the overall state judgments are used by the model. Thus considerable information is sacrificed--as in any forecasting method.

- Since the Bayesian revision of model parameters assumes that the true parameters are constant (satisfactory), states must be chosen and defined with great care. If the states are not or cannot be defined such that transition probabilities and waiting times are stable, the Bayesian revision process will be inappropriate, and the model will frequently require direct reassessment of parameters. It is not clear that, with a small number of states, any reasonable degree of stationarity can be achieved by the present implementation.

- In many cases of application, the real interest will be on the forecasting of crises or other relatively unlikely states. The usefulness of any kind of Bayesian revision based upon the states alone is questionable since empirical state-data will rarely be encountered in the "crisis" range and because prior assessments will be strongly opposed to the occurrence of crises. This is simply a case of not being able to collect an empirical data base on rare events; most of the Bayesian revision will improve estimates of parameters related to relatively normal and uninteresting states.

To illustrate some of these points, consider an example (Job and Duncan, 1980). Exhibit 1 shows a 30-day forecast from 11 March 1978 concerning a Middle East situation, on a day when the analyst judged the current state to be 2. On 15 March (day 4 of the forecast in Exhibit 1) the analyst judged the state to be 4, an event to which his 11 March forecast assigned a 2.6% likelihood. The crisis was thus not "forecast" in any sense. Exhibit 2 shows the forecast made on 15 March, and the fact that the near-term likelihood of war (state 5) has greatly increased. This increase, of course, is due primarily to the observed change to state 4, while any previous Bayesian revisions of parameters amount to fine tuning. While limited in their ability to forecast crises in the short term, present implementations may be of far greater value as a long-term forecasting methodology. They appear also to be more appropriate for forecasting underlying conditions than crises or events, and perhaps could be used in conjunction with an inference model to help crisis forecasting. The actual use of the methodology depends strongly on the states that are used; the present choices do not appear to use the full power of the methodology, and we discuss later some alternate choices.

2.2.3 Effectiveness of current assessment procedures. In actual use, the Markov renewal model must often be modified by assessing new prior distributions for waiting times and the transition matrix. This is necessary because of the fundamental non-stationarity of those parameters. Tests conducted to date indicate such modifications must be made relatively often. This means that there is relatively little Bayesian revision of parameter estimates, and that the model forecasts depend heavily upon the assessed prior distributions.

The parameters of the waiting time distribution are--at least in some implementations--assessed according to a fractile assessment procedure. Such procedures are known to produce biased assessments in that the fractiles tend to be too close to the median of the distribution and, thus, the distribution is too tight

EXHIBIT 1

Probability Forecast

Israeli-Syrian Interaction Conditions

30 day Horizon, Dated 11 March 1978

(Reproduced from Job and Duncan, 1980)

Days in Future	Interaction Condition				
	1	2	3	4	5
1	3.3	92.0	3.6	0.68	0.46
2	6.2	85.0	6.8	1.3	0.91
3	8.7	78.0	9.7	1.9	1.4
4	11.0	73.0	12.0	2.6	1.9
5	12.0	68.0	14.0	3.2	2.4
6	14.0	64.0	16.0	3.9	2.9
7	14.0	60.0	17.0	4.5	3.5
8	14.0	58.0	18.0	5.2	4.0
9	14.0	56.0	19.0	5.8	4.6
10	14.0	55.0	19.0	6.4	5.2
11	14.0	54.0	19.0	7.0	5.8
12	13.0	54.0	19.0	7.5	6.3
13	13.0	53.0	19.0	8.0	6.9
14	12.0	53.0	19.0	8.4	7.4
15	12.0	52.0	19.0	8.8	7.9
16	11.0	52.0	19.0	9.1	8.5
17	11.0	52.0	19.0	9.4	8.9
18	11.0	52.0	19.0	9.7	9.4
19	11.0	51.0	19.0	10.0	9.8
20	10.0	51.0	18.0	10.0	10.0
21	10.0	50.0	18.0	10.0	11.0
22	10.0	50.0	18.0	11.0	11.0
23	10.0	49.0	19.0	11.0	11.0
24	10.0	48.0	10.0	11.0	12.0
25	10.0	48.0	19.0	11.0	12.0
26	10.0	47.0	19.0	11.0	12.0
27	10.0	47.0	19.0	11.0	13.0
28	11.0	46.0	19.0	12.0	13.0
29	11.0	46.0	19.0	12.0	13.0
30	11.0	45.0	19.0	12.0	13.0

EXHIBIT 2

Probability Forecast
Israeli-Syrian Interaction Conditions
30 day Horizon, Dated 11 March 1978
(Reproduced from Job and Duncan, 1980)

Days in Future	Interaction Condition 1	2	3	4	5
1	0.000014	0.083	2.1	95.0	3.0
2	0.00049	0.28	4.9	87.0	7.4
3	0.0031	0.59	7.5	80.0	12.0
4	0.011	1.0	9.9	72.0	17.0
5	0.029	1.5	12.0	66.0	21.0
6	0.065	2.2	13.0	60.0	25.0
7	0.13	3.1	14.0	55.0	28.0
8	0.24	4.2	15.0	50.0	31.0
9	0.4	5.4	15.0	46.0	33.0
10	0.62	6.9	15.0	43.0	34.0
11	0.91	8.5	16.0	40.0	35.0
12	1.3	10.0	16.0	37.0	36.0
13	1.7	12.0	16.0	35.0	36.0
14	2.1	14.0	16.0	32.0	36.0
15	2.6	15.0	16.0	30.0	35.0
16	3.0	17.0	17.0	29.0	34.0
17	3.5	19.0	17.0	27.0	34.0
18	4.0	20.0	18.0	26.0	32.0
19	4.5	22.0	18.0	24.0	31.0
20	4.9	23.0	19.0	23.0	30.0
21	5.3	25.0	19.0	22.0	29.0
22	5.7	26.0	20.0	21.0	28.0
23	6.1	27.0	20.0	20.0	26.0
24	6.4	28.0	21.0	19.0	25.0
25	6.8	30.0	21.0	19.0	24.0
26	7.0	31.0	21.0	18.0	23.0
27	7.4	32.0	21.0	18.0	22.0
28	7.6	33.0	21.0	17.0	21.0
29	7.9	34.0	22.0	17.0	20.0
30	8.1	35.0	22.0	16.0	20.0

(Lichtenstein, Fischhoff, and Phillips, 1977). Suggested solutions to this difficulty are discussed later.

3.0 THE POTENTIAL VALUE OF MARKOV RENEWAL MODELS

In this section we discuss whether, and in what ways, using a Markov renewal process improves on other, simpler, stochastic processes for international forecasting. Duncan and Job, (1977, p. 12), arguing for the semi-Markov process, write that the discrete time used in other Markov models is inappropriate, because assuming that movements are made at regular intervals "almost certainly is not true of movements in international interaction processes". While this is undoubtedly true, the observation does not mean that a discrete-time model is inappropriate. Indeed, in the applications thus far, the forecasts have been made at daily intervals, showing that a discrete time model, which may be viewed as taking snapshots of a continuous time process, may well be appropriate. To discuss this further, we look at the available stochastic models and look at their available flexibility.

3.1 Semi-Markov Process in Continuous Time

This is the model used by Duncan and Job. It has the flexibility of assessing transition probabilities between states, and also general distributions of waiting times for a transition which may vary both by the state before, and the state after, the transition. To simulate this, we visualize that after a transition, a random choice is made, according to the probabilities of the relevant row of $\underline{J}$, determining where the next jump will be, and that we then sample from the relevant element of $\underline{\underline{F}}$ to discover when that transition will be made.

3.2 Semi-Markov Process in Discrete Time

Such a process appears not to have been studied in the literature, but it is simply a discretization of the continuous time case and is, in effect, what Duncan and Job use. It probably has little advantage in computational ease, and there seems little reason to change the current model in its favor.

3.3 Markov Process in Continuous Time

This differs from type 3.1 in that the waiting time distributions are independent of the states to which a jump is made, i.e., the distribution depends only on the current state. Such a model thus loses some of the flexibility of 3.1. However, we are unconvinced that in the type of applications to international forecasts that have been carried out this is an important loss. To understand this, imagine that we have several years of data concerning the transitions between states. Then the extra dimension of 3.1 means looking only at the transitions that actually occurred from the distribution of times in i among all the actual transitions from i to k. Certainly we may expect a substantial difference in the <u>number</u> of each type of transition, but it is unclear that the waiting distributions will differ. Indeed, if they did, we could use bayes theorem to make inferences about which transition has become relatively more likely to come next, given an observed sequence of days in state i. For example, if the expected wait to state j is 7 days, and to state k is 30 days, and we have been in state i for 20 days, the transition to state k now appears relatively more likely than indicated by $\underline{J}$. It is not clear that such an inference is in fact valid, or at least that the assessments necessary of an analyst can be meaningfully obtained. As an example, consider the case study in Job and Duncan (1980). Here the expected waiting time for the transition $2 \rightarrow 3$ is 10 days, and for $2 \rightarrow 5$, 20 days. From this data, were we to observe state 2 for over 70 days (as in fact occurred), then our model implies that we are much more likely to move into war than if state 2 had only been observed for 2 days. We find it hard to believe that, a priori, the analyst felt this to be true. If the waiting time is assumed to be identically distributed for all transitions out of i, the deduction would not be made. Thus we suspect this simpler model may in some cases be more appropriate than type 3.1, but such a suspicion should be checked in a field test.

3.4 Markov Process in Continuous Time with Poisson Process

This is the same model as 3.3, except that the waiting time distributions are now constrained to be of exponential form, which is equivalent to saying that there is a Poisson process, i.e., that the "lack of memory" property holds true, and that to calculate the distribution of the waiting time from a time t, all that needs to be known is the state X(t), but not how long the system has been in that state. Such an assumption may or may not be reasonable in a given case. This assumption is further discussed in the next section. For now, we note that there may be little loss in flexibility in using 3.4 rather than 3.3, since Duncan and Job do not exploit the full possibilities with regard to waiting time distributions but stick with log-normal distributions.

3.5 Markov Chain--Markov Process in Discrete Time

This is the usual form used in Markov modelling, where a transition matrix is used to calculate changes through time step by step. The transition matrix may vary through time, but often is assumed constant. This latter is the discretization of model 3.4 and, in day-to-day forecasting, would produce very similar results to model 3.4. Note that the "lack of memory" property is true with this model.

4.0 THE TREATMENT OF TIME IN FORECASTING INTERNATIONAL AFFAIRS

There are two distinct ways in which time enters the problems of forecasting international affairs. First, on a given day, we can make a forecast about the future. Second, we can change that forecast about the future on a different day, in the light of new information. Duncan and Job deal with the first time dimension by using a Markov renewal process, and with the second type by using Bayesian updating on the parameters of that process. In this section we examine these separate aspects of the time characterization problem.

4.1 A Priori Forecasts

Note first that not only do we have the option of choosing the model, but also of choosing the state space. There is no theoretical restiiction on the number or type of states. For example, a perfectly valid state might be "Country A is in fairly high tension, has been for 3 days, was in low tension for 10 days before that, has just changed government, and the weather is 2^o colder than average for this time of year". In the extreme, the state at time t could be the total history of the world to time t. In general it is true that the more complex the state-space description, the simpler the stochastic model will be. In effect, increasing the complexity of the state-space means that more of our knowledge about the real-world system is captured in it, and thus that less needs to be captured by the stochastic model. However, a complex state-space can mean an inordinate amount of assessments, so using a simpler state-space, and making fewer general statements about the world which encapsulate our knowledge, leading to a more complex stochastic model is necessary. The problem, and where the real modelling skill is necessary, is to choose a state-space that is sufficiently simple to be manageable, but rich enough for the general statements that are made to be meaningful.

An obvious example concerns the problem of stationarity. This concerns the criticisms often voiced against Markov modelling, that they often assume transition probabilities to be constant over time. This is clearly false if the underlying process changes; for example, if a new government enters power. However, if the state space included a description of the government, then this non-stationarity would not arise.

This type of observation has been used to argue that more complex state-spaces are required. However, the more states, the harder both assessments and interpretations of results become. If we are attempting to forecast crises, there is so much detailed knowledge the analyst has about mechanisms etc., that a state-space model, to be meaningful, would have to be huge. We do not believe that such a stochastic model would be of value. Rather, to forecast crises, an inference model should be built to capture the analysts' beliefs. Then one of the key inputs to such a model would be the "states" as used by Duncan and Job. Thus we might envisage a hierarchical inference (H.I.) model, whose output was P(crisis in next week), conditioned on the level of stability of the country. Then we could use Duncan and Job's model as shown in Exhibit 3. (Note that we still are not looking at how our forecasts vary as real-time progresses; the exhibit refers only to forecasts made on one particular day.) Here we use the Markov renewal process to derive probability distributions over the states, and then use these "state

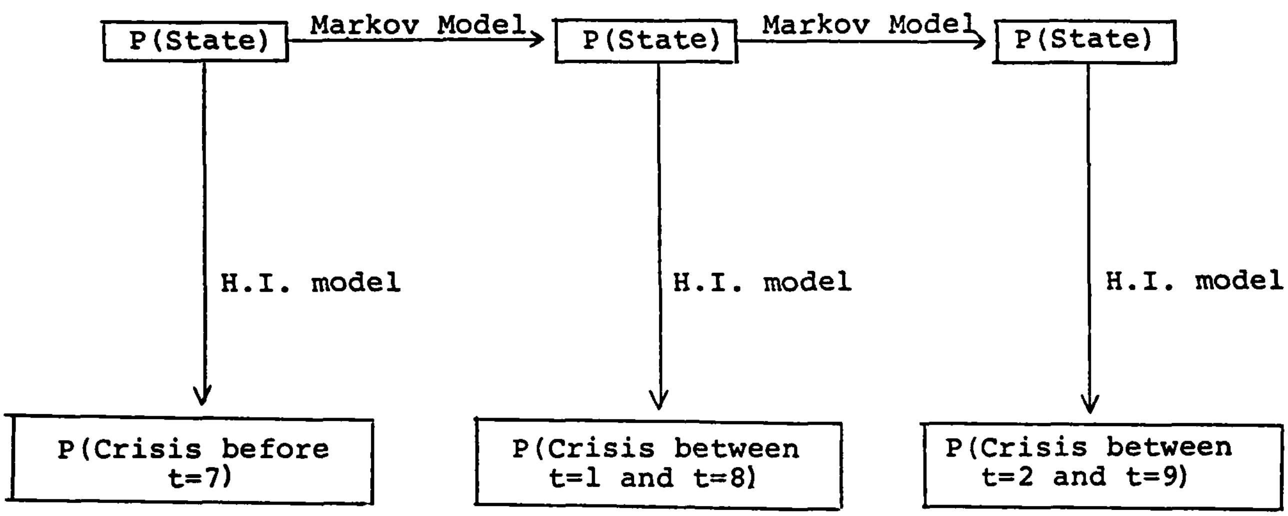

EXHIBIT 3

INCORPORATION OF HIERARCHICAL INFERENCE IN A MARKOV RENEWAL MODEL forecasts" in our H. I. model to forecast crises. We may still need to include, say, a change of government in our state space description, but since our Markov model is at a very general level, the number of states required should remain small.

It should be noted that present implementations define states on a continuum of stability, but that in general states can be qualitatively, as well as quantitatively, different. When such a continuum is used, it may be valid to argue that, in continuous time, one can only move to a neighboring state, e.g., tension will escalate to war, rather than jump there, so the probability of a transition directly from 3 to 5 is zero. Of course the time in an intervening state may be minimal, but the transition is there. This is perhaps more intuitive than allowing any direct jump An intriguing possibility is that we might view the process as one with continous states (level of tension) and look at the resulting Markov process. Such an approach has not been followed through here, but it might prove to be of value.

In the light, then, of the schema presented in Exhibit 3, we may look to see which type of model discussed in the previous section will be of use. Since Duncan and Job have developed the model and computer program for 3.1, we do not include in our evaluation the factor of computational complexity, so, although discrete time processes might be easier computationally, we shall assume we shall use the continuous time model they have built, but possibly with fewer, or simpler, assessments, and maybe not utilizing its full potential power.

The first question concerns the use, or otherwise, of waiting-time distributions conditioned on the next transition. As argued in the previous section, in most cases we consider this richness to be of little value. For example, it is unlikely that if a representation of the process shows a country to be in a stable state for a certain number of days our prior probability of where the next transition will be is affected. On the other hand if the representation is in a war state, then knowing that the war state has obtained for 20 days may lead us to believe that a certain transition is now more likely than it was before. In either case, by asking the question in this form we can decide if the differences are relevant. We may also, by exploiting Bayes' theorem, ask questions about the shift in probabilities to assess the differences in waiting time distributions in a manner that is more intuitive to an assessor than is a direct elicitation.

The second question concerns the use of a Poisson process rather than a more general one. This can again be decided by asking the analyst whether, and in what ways, knowledge of the current length of stay of a representation in a particular state would affect his/her expectation of a jump. We would anticipate that from a stable state, the "lack of memory" property may be satisfied, whereas from a war state, the situation may again be more complex. For example, the analyst may believe that a war could not last for more than twenty days, due to supply problems, in which case he/she would expect a quicker end to a war after 19 days than initially. In such a case, the log-normal distribution currently used may be more appropriate. It should also be noted, however, that "war" may be an inappropriate state to have in the stochastic model when we operate under the schema of Exhibit 3, as it probably would be viewed as one of the variables of interest of the H. I. model, rather than a low-level input. If such proves to be the case, both the independence of waiting time and next state, and the lack of memory property, may prove to hold for all states. Note that different analysts may provide different answers to these questions, in which case a Poisson process may properly model one set of beliefs, but not another.

4.2 Updating Forecasts

We now address the issue of the second dimension of time; of how we should alter our forecast in the light of new information. Duncan and Job add a Bayesian updating model to the Markov renewal process. Having made a forecast on day 1, the analyst then needs only to decide upon the current state, on day 2, in order to generate a new forecast on that day.

By allowing the analyst to judge only the current state, we allow him very little latitude to express his opinions about the new information. If all of this information is not captured in the state description, we are surely losing something. Of course, one could overcome the problem by using a very detailed state-space, but as we have argued above, this is probably undesirable in terms of ease of elicitation. Rather, we believe that the new information will result not only in possible changes of state, but also in possible changes of the parameters of the model. For example, an observed change of government may cause an analyst to change his $\underline{\underline{F}}$ and $\underline{\underline{J}}$ matrices. Such a shift could not be effected by Bayesian updating, unless we had elicited probabilities of receiving that type of information, conditional on various future states of the world, from the analyst, in advance. Such an assessment effort would be foolish and very costly.

The Bayesian updating incorporated by Duncan and Job in their model does not address this type of probability shift, but rather one resulting from observed transitions between states. The difficulty with this is that one of the shifts of probabilities discussed above will probably occur before very many transitions, so that the priors will be very little altered before the analyst needs manually to alter F and J. In practical use of Duncan and Job's model this has often proven to be the case. After a small number of transitions, a fundamental change occurs, forcing a manual change in F and J. It should also be noted that if we are operating under the schema of Exhibit 3, the new information may also change the states that are considered to be appropriate for the model, and may also change the structure of, or inputs to, the H. I. model.

The difficulty we are discussing is another problem of non-stationarity, but a different one to that discussed in the previous section. Previously we were querying whether an analyst would have a prior belief that and how the underlying process would change, and we believe that that question could often be answered in the negative, since the uncertainty about changes in the underlying process could be captured in the probabilities. We are now querying whether an analyst may have a posterior belief in a change, after he has received information, and the answer surely is in the affirmative. If he has information stating that the change has occurred, he would like this to be modelled. The point is that if this type of change is not permitted, we are assuming that our initial modelling effort modelled not only the analyst's beliefs at the time of modelling, but also all his possible beliefs over the future, and his reactions to any information that he might receive in the future. Equally, we would be assuming that all future information would be irrelevant except in so far as it is captured by the designation of "current state". We believe these assumptions to be unrealistically strong. This is true not only of Markov renewal models, but also of any forecasting model that updates automatically. Without an inordinate amount of structure elicitation, and assessments that attempt to cover every possible contingency, there will always be a need for the analyst to interact with the model as real-time progresses.

5.0 SUGGESTED MODIFICATIONS AND RESEARCH

Based on the discussion in the previous sections, a number of recommended paths for further modifications and research can be identified. Three general paths are discussed in this concluding section.

5.1 Inclusion of Indicator-Based Inference Models

In Section 3.0, the need for the use of indicators was developed. The related problems of the coarsely defined states and the nonstationarity of the transition and waiting time parameters lead to this need. Properly developed, inference stages based on indicators could be incorporated in current models; such changes should lead to improved forecasts through better use of available information and should enable the construction of more responsive models which are relatively free of nonstationarity problems. Exhibit 3 represents one path this research might follow.

5.2 Modification of Parameter Assessment Methods

The present implementation of the Markov renewal model is primarily based upon subjective assessments of model parameters. Because the rapid changes which characterize international affairs preclude extensive Bayesian revision of these subjective assessments, due to the need to directly revise parameters when the situation changes, these "prior" assessments largely determine the forecasts produced. Because of this, one of the ways in which the model and the forecasts it produces can be improved is by improving the assessment of prior distributions. DARPA has supported much research (e.g. Edwards and Seaver, 1976) into how to elicit subjective probabilities, and some of these procedures could be incorporated into the assessments required for the Markov renewal model. Procedures other than the fractile procedure currently used--which has drawbacks discussed in Section 2.2.3--could be employed to obtain better assessments of the waiting time distribution. Previous DARPA-supported research has indicated at least one procedure which is less likely to produce biased assessments (Fuji, Seaver, and Edwards, 1977; Seaver, von Winterfeldt, and Edwards, 1975). The assessor can be given a few waiting times and asked for the probability that the waiting time will be less than the time given. These probabilities can then be used to specify the waiting time distribution through a procedure similar to that being used in the current model. Such a change would require only minor software modifications.

The elicitation of the Dirichlet prior distributions for the transition matrix could also be improved. Several alternative procedures are available. The equivalent prior sample (EPS) procedure asks the assessor what sample from the relevant process would produce a state of knowledge in the analyst equivalent to his/her current opinion. The hypothetical future sample (HFS) procedure determines the parameter $\underline{n}$ (the number of points referred to above) by comparing an initial assessment with one made given a hypothetical additional sample is known to the assessor. Winkler (1967) describes these procedures and an empirical test of their use. The parameters of the Dirichlet distribution could also be obtained by using the fractile procedure or the alternative described above to assess the marginal distribution of the parameters, each of which is a beta distribution.

5.3 Evaluation of Model Forecasts

Most of the preceding discussion of the current Markov renewal model implementation has necessarily been limited to indirect arguments. A more satisfying evaluation would, of course, be based upon a careful empirical test. The model has value largely to the extent that it can improve the ability of analysts to anticipate changes in the international situation. Whether or not it performs that function must be decided on the basis of empirical data.

Several relatively inexpensive evaluation methods are possible. An actual comparison of aided vs. unaided analysts' forecasts could be made. The scoring of forecasts could be accomplished along the lines of scoring rule procedures for probability assessments (e.g. Murphy and Winkler, 1970). Several variants are possible in the design of such an experiment.

In addition, since states are being forecast and observed states are presently supplied daily to the model, it would be quite simple to score the model's performance in an ongoing fashion. While fairly large numbers of trials would be required to obtain stable estimates of scores, this "self-evaluation" feature could prove useful in multiple ways: e.g., it could provide automatic warning to an analyst at times when changes render the model ineffectual, and it could be used to compare competing variants of the methodology.

Another form of empirically-based evaluation would examine specific assumptions of the model. While not as conclusive as the preceding evaluation methods, these kinds of tests could probably be performed now on the basis of previously collected data. Examples include testing the validity of the Markov assumption and comparing assessed distributions--including those assessed under a variety of methods--with actual occurrences.

5.4 Modification of Model Output Formats

Since the Markov renewal model conceptualizes the international affairs process as a sequence of state transitions, and assumes that analysts can assess parameters of such a process, alternate formats for model output should probably be considered (if they have not already been rejected for cause). In addition to the probability that a system will be in a particular state at a particular time, it might help the analysts' understanding to provide them with estimates such as the probability that a particular state will be reached by a particular time, or the expected time for a particular state to be reached. All of these calculations could be performed on the basis of the present model and software.

6.0 REFERENCES

Duncan, G. T., and Job, B. L. Crisis forecasting using the Markov renewal model (Technical Report 131). Pittsburgh, PA: Department of Statistics, Carnegie-Mellon University, July 1977.

Duncan, G. T. Bayesian statistical estimation in the finite state Markov renewal process (Technical Report 139). Pittsburgh, PA: Department of Statistics, Carnegie-Mellon University, December 1977.

Edwards, W., and Seaver, D. A. Research on the technology of inference and decision (SSRI Research Report 76-7). Los Angeles: University of Southern California, Social Science Research Institute, October 1976.

Fujii, T., Seaver, D. A., and Edwards, W. New and old biases in subjective probability distributions: Do they exist and are they affected by elicitation procedures? (SSRI Research Report 77-4). Los Angeles: University of Southern California, Social Science Research Institute, August 1977.

Job, B. L. Signals, triggers, and thresholds: The problem of analyzing turning points in international conflict processes. Paper presented at the American Political Science Association Annual Meetings, Washington, D.C., 1977.

Job, B. L., and Duncan, G. T. Probability forecasting in international affairs. Paper presented to 1980 meetings of the International Studies Association, Los Angeles.

Lichtenstein, S., Fischhoff, B., and Phillips, L. D. Calibration of probabilities: The state of the art. In H. Jungerman and G. de Zeeuw (Eds.), Decision making and change in human affairs. Amsterdam: D. Reidel, 1977.

Murphy, A. H., and Winkler, R. L. Scoring rules in probability assessment and evaluation. Acta Psychologica, 1970, 34, 273-286.

Seaver, D. A., von Winterfeldt, D., and Edwards, W. Eliciting subjective probability distributions on continuous variables (SSRI Research Report 75-8). Los Angeles: University of Southern California, Social Science Research Institute, August 1975.

Winkler, R. L. The assessment of prior distributions in Bayesian analysis. Journal of the American Statistical Association, 1967, 62, 776-800.

DISTRIBUTION LIST

Director
Defense Advanced Research Projects Agency
Attention: Program Management Office
1400 Wilson Boulevard
Arlington, Virginia 22209

Defense Technical Information Center
Attention: DDC-TC
Cameron Station
Alexandria, Virginia 22314

DCASMA Baltimore Office
Attention: Mrs. Betty L. Driskill
300 East Joppa Road
Towson, Maryland 21204

Decisions and Designs, Inc.
Suite 600, 8400 Westpark Drive
P. O. Box 907
McLean, Virginia 22101

Director
Defense Advanced Research Projects Agency
Defense Sciences Office
Cybernetics Technology Division
Attention: Dr. Judith Ayres Daly
1400 Wilson Boulevard
Arlington, Virginia 22209